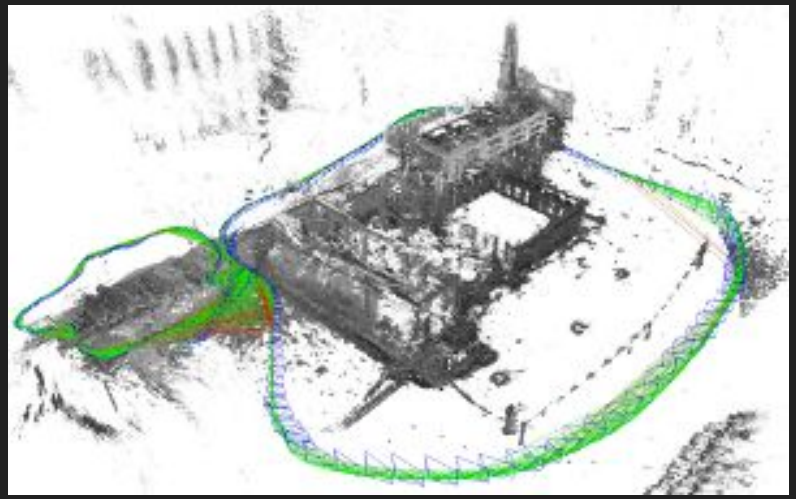
Machine Learning for Robotic Vision

Tom Drummond

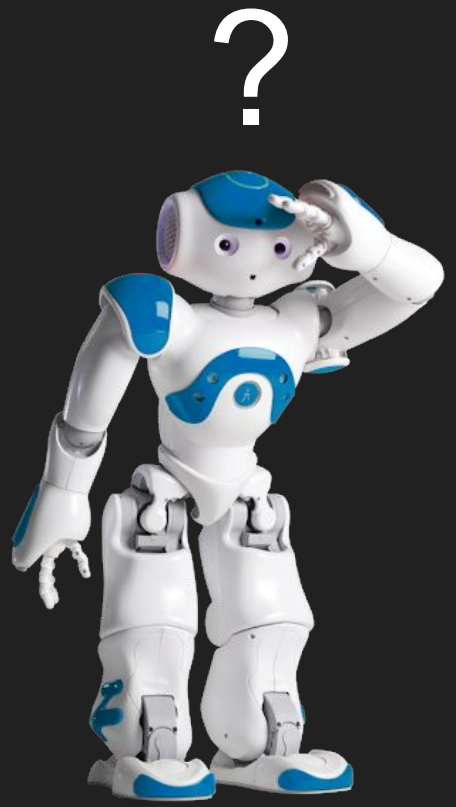
Robots need to model the world

Where things are (Geometry)



What things are (Semantics)

Robots need to know what they don’t know

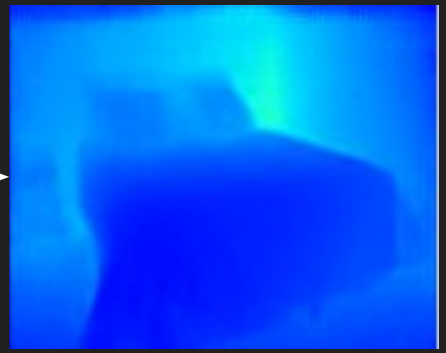


“There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns – the ones we don't know we don't know.”

- Donald Rumsfeld

Deep learning can estimate depth (Geometry) from RGB

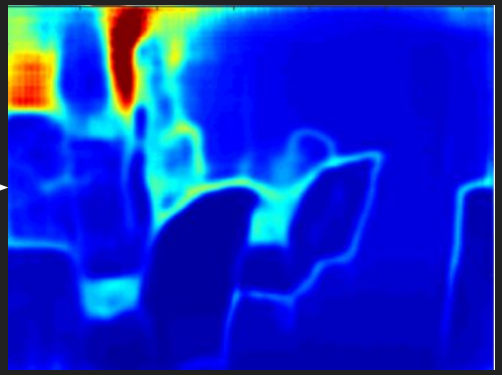
Depth estimation network



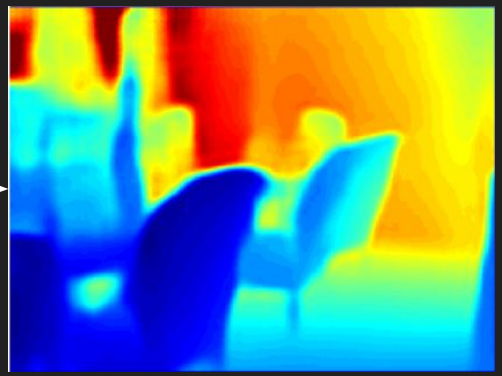
Can minimize L_2 loss to ground truth

Better to learn uncertainty as well as depth estimates

Depth estimation network

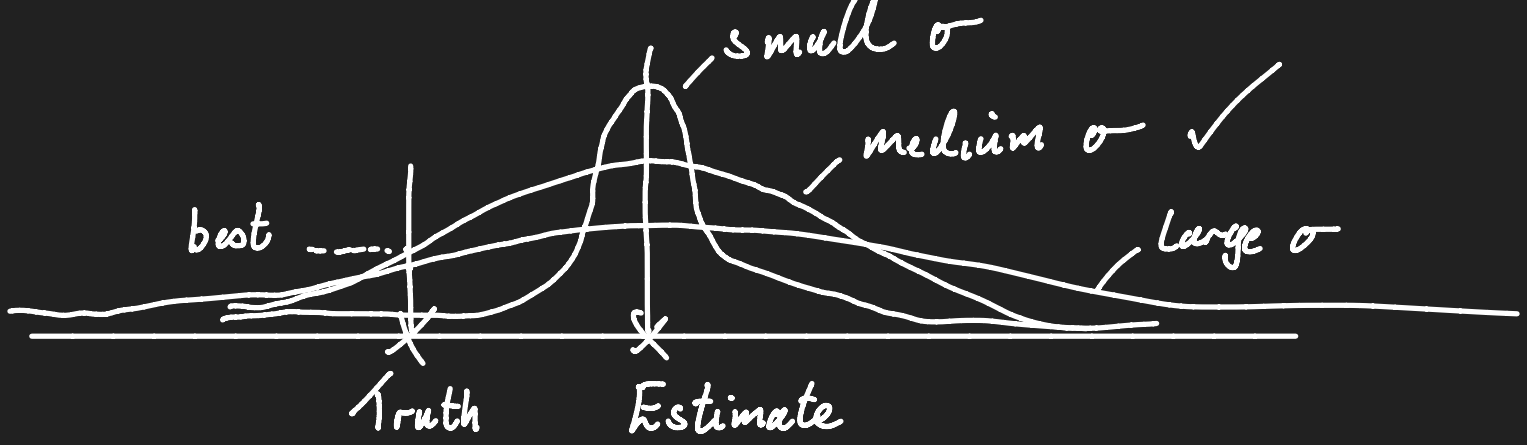


Depth (μ)

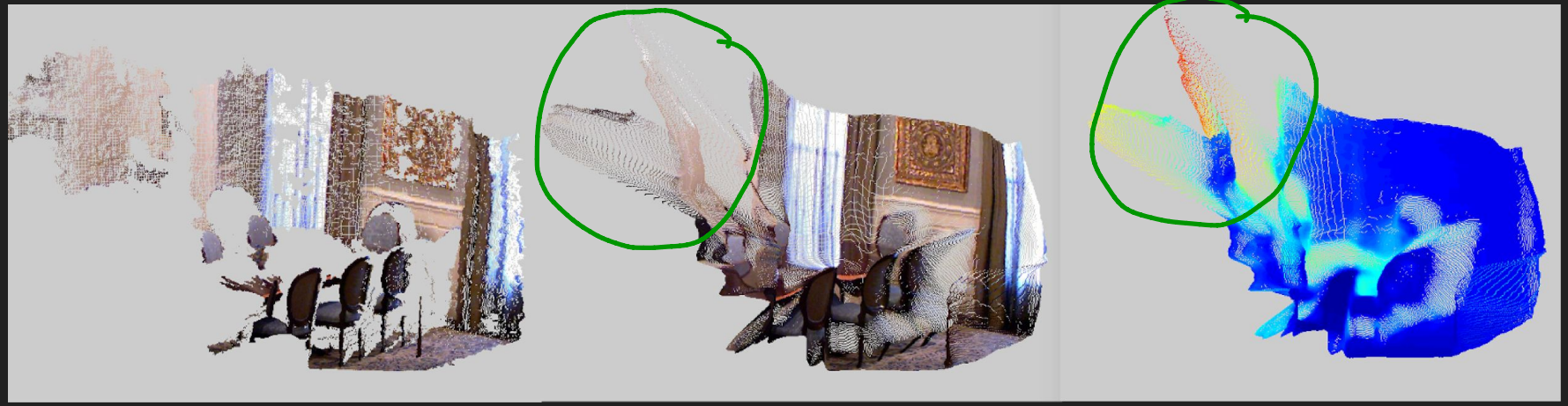


Uncertainty (σ)

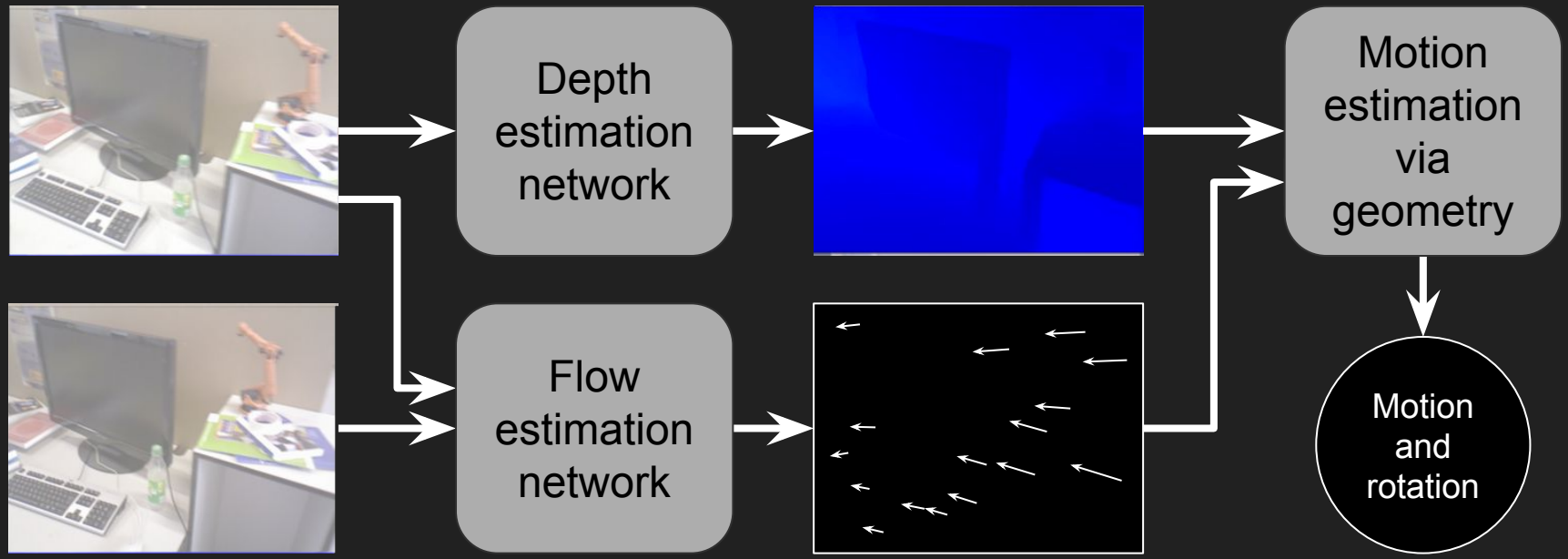
Mean (μ) and sigma (σ) imply a pdf
now minimize log loss to ground truth


small σ
medium σ ✓
best
large σ
Truth
Estimate

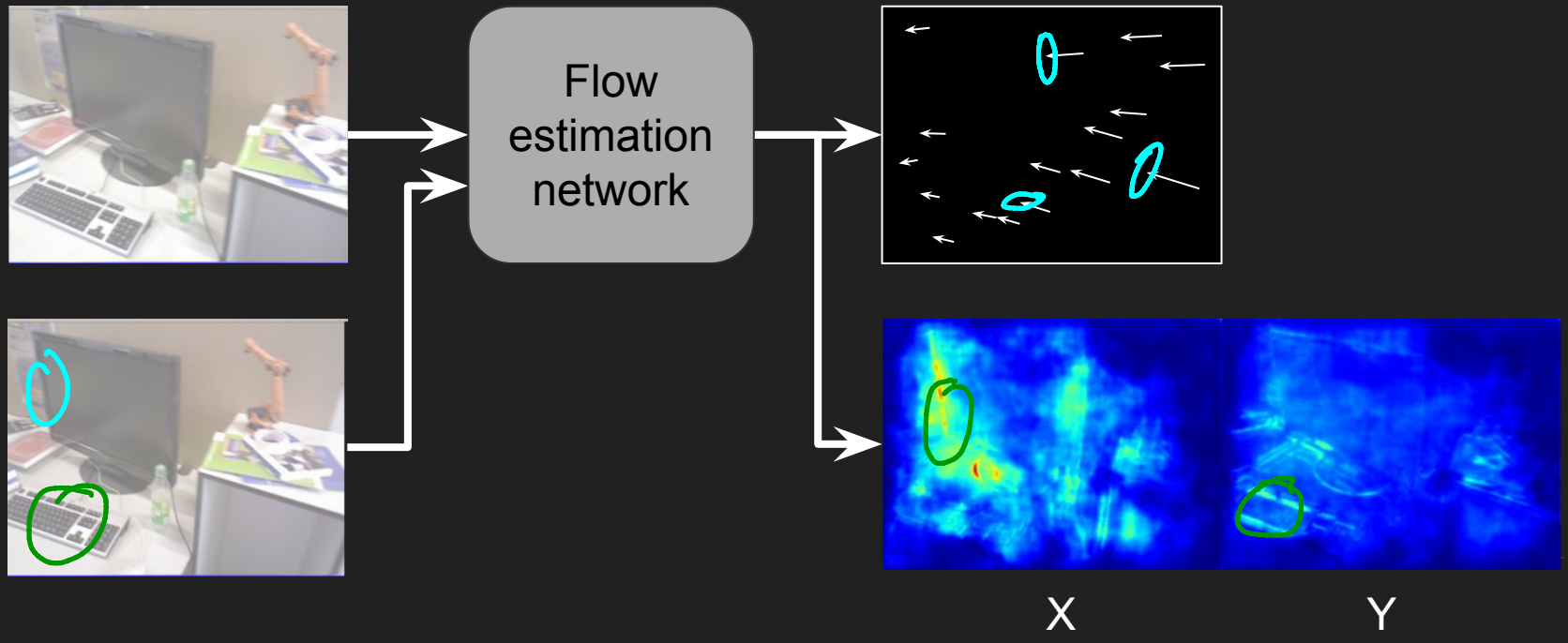
Regions of high uncertainty correspond to large errors



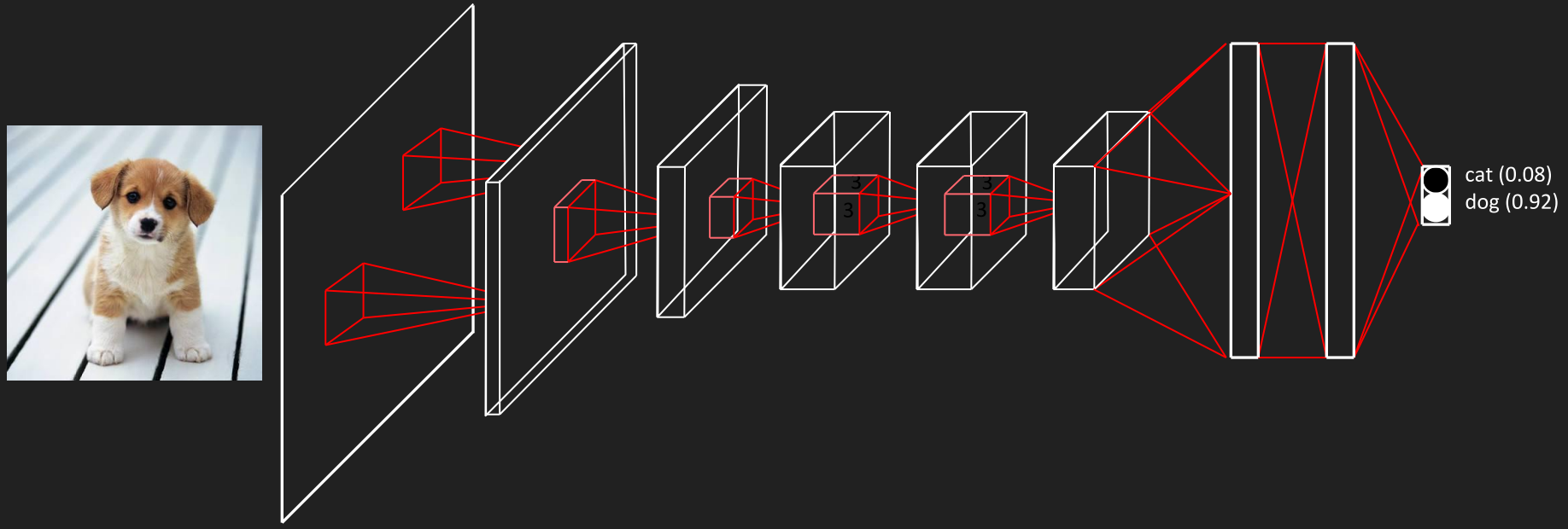
Can build this up into a Visual Odometry pipeline



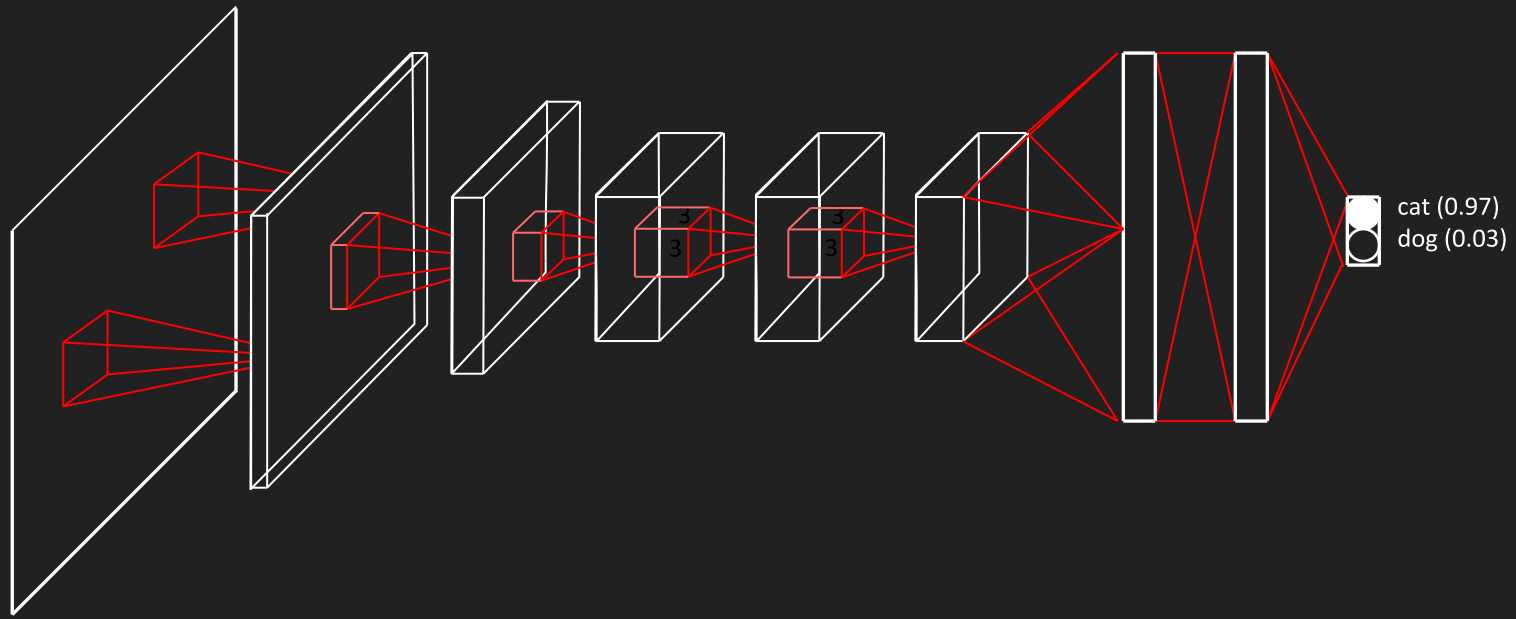
Want the network to estimate confidence (Information)



Recognising objects in the environment



Recognising objects in the environment



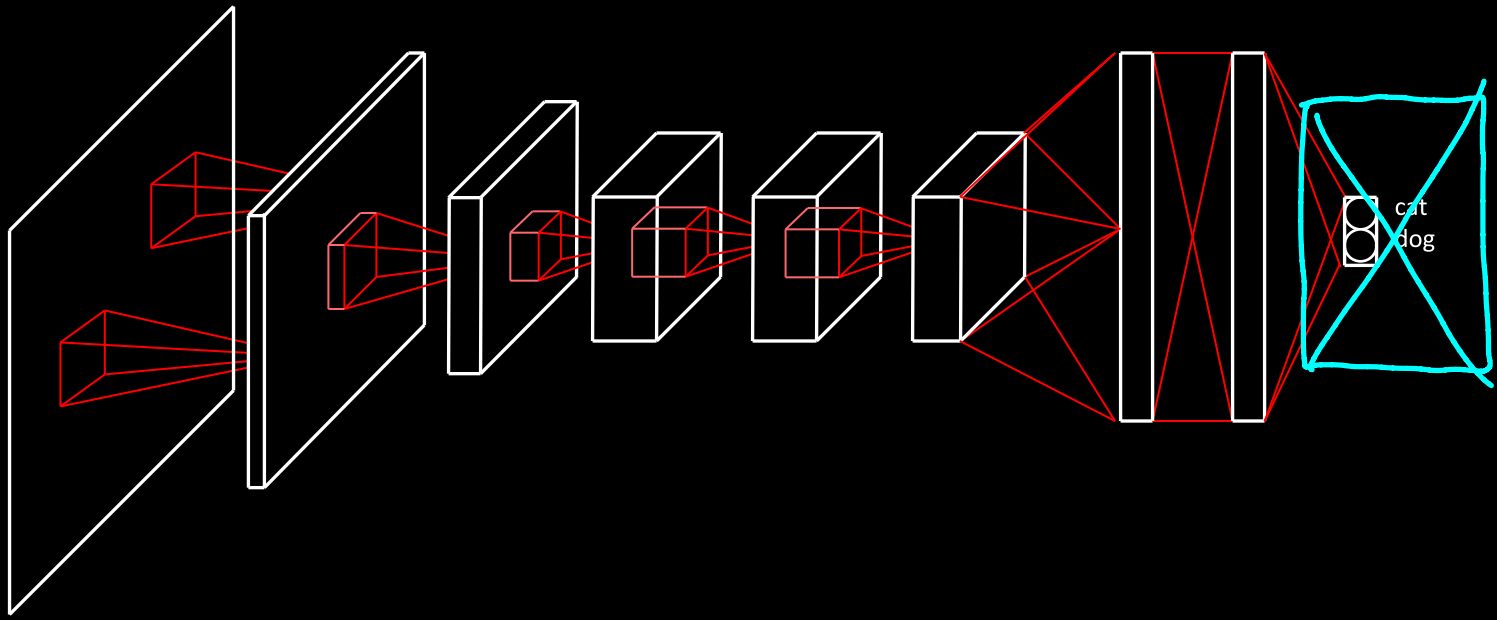
What happens with photoshopped chimeras?

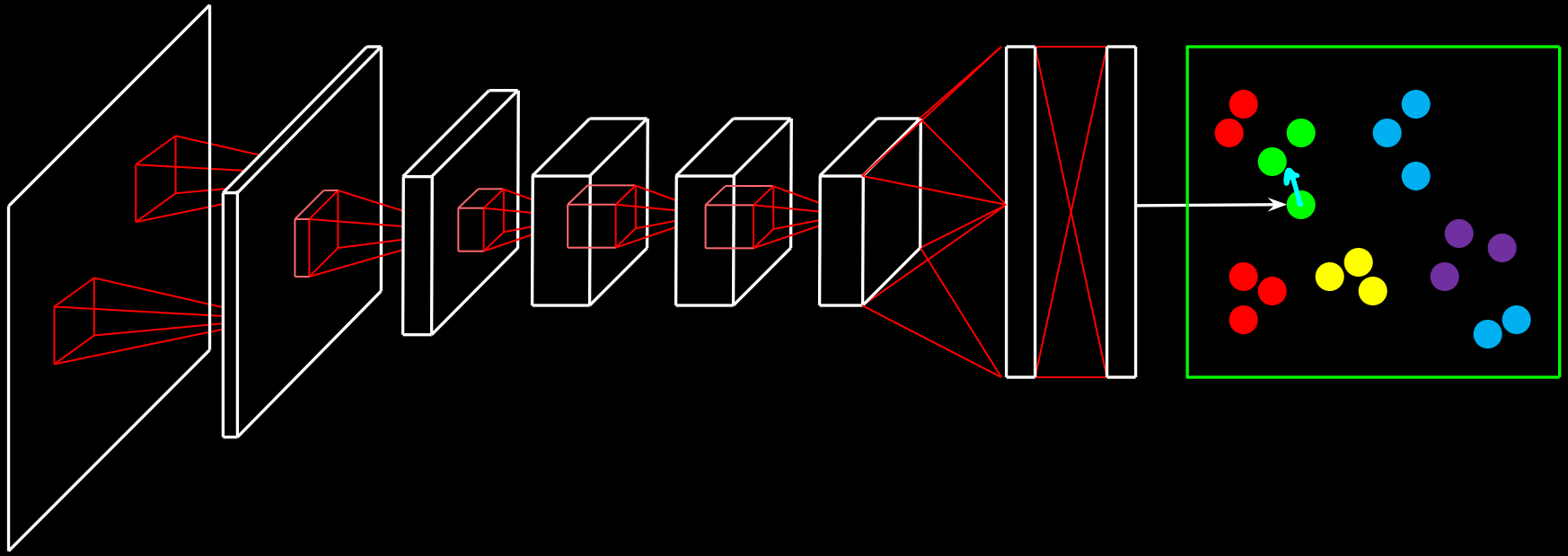
cat (0.4)
dog (0.6)

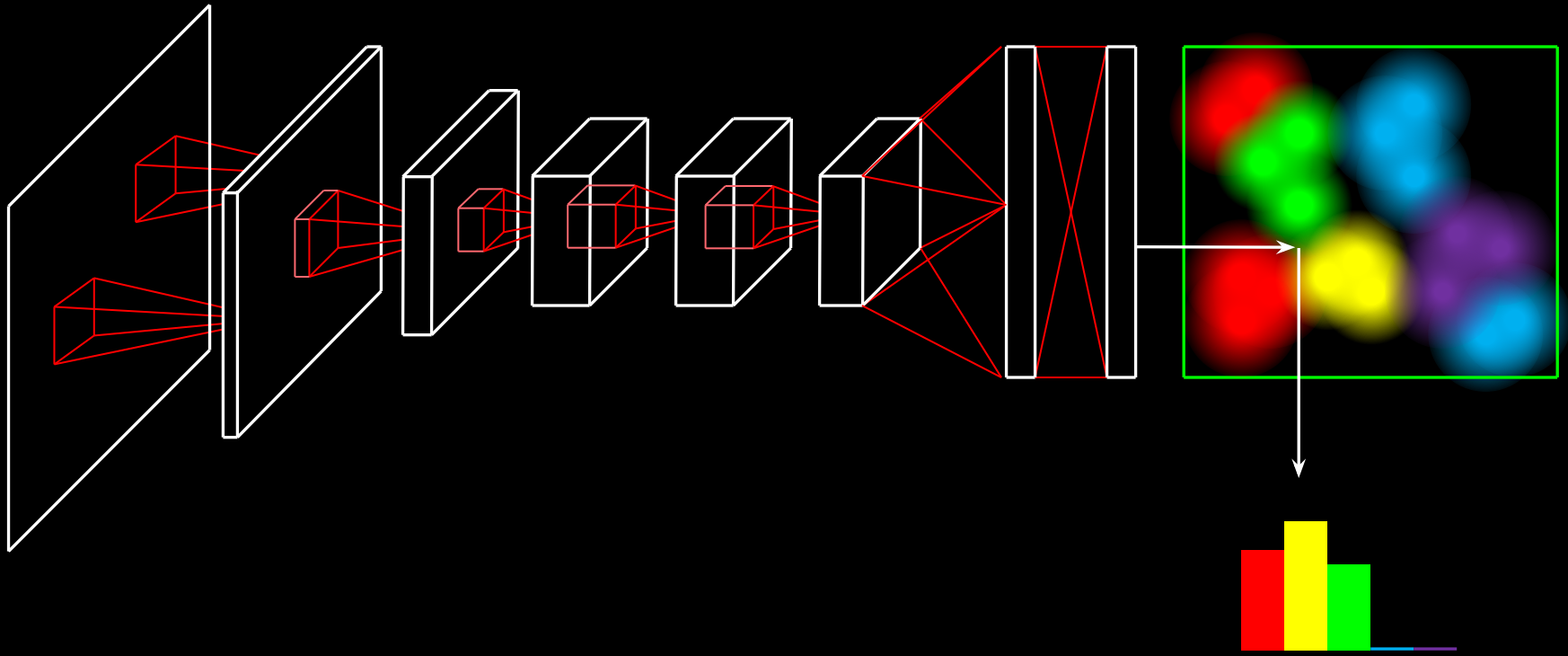
cat (0.7)
dog (0.3)

Or something completely different?

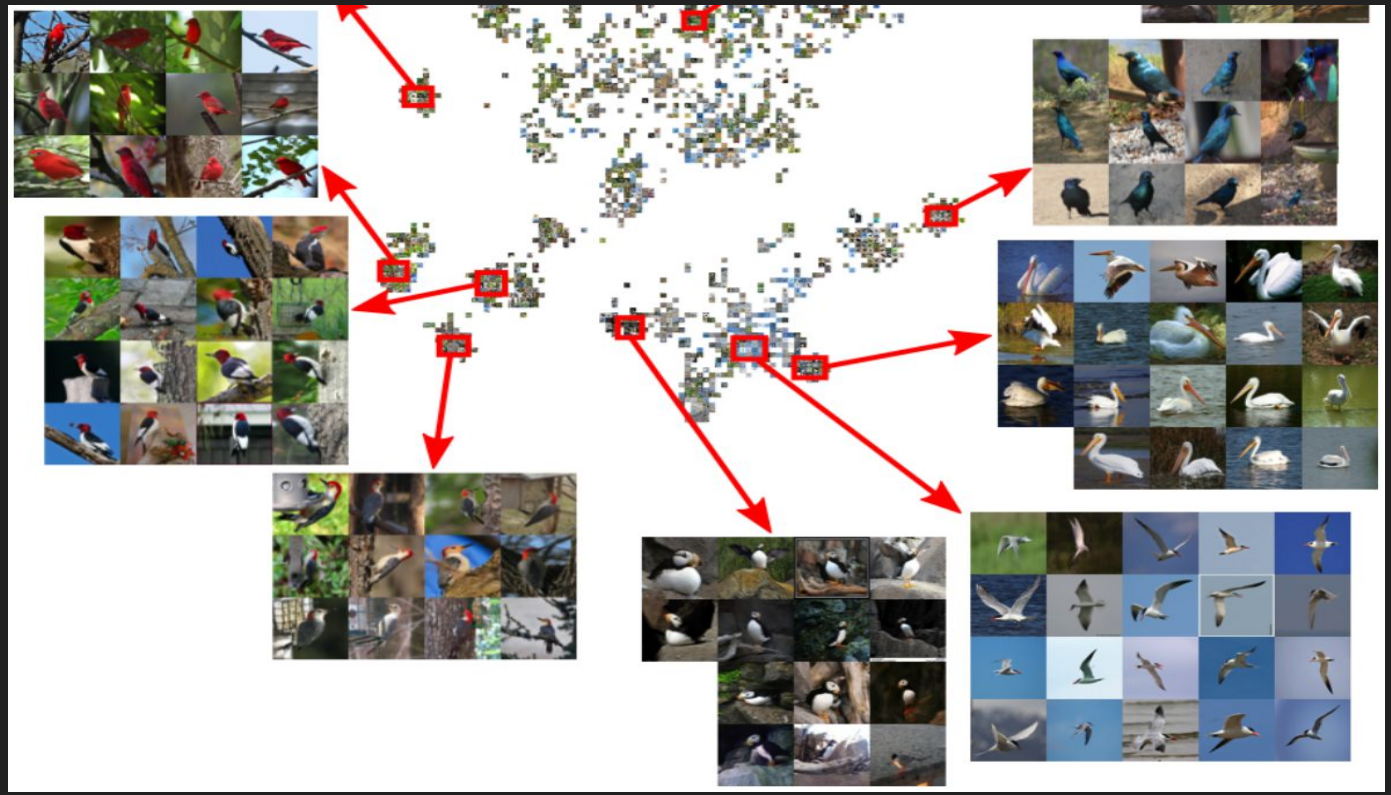
cat (0.99)
dog (0.01)


cat
dog





Learned mappings can be transferred to new classes



Learned mappings can be transferred to new classes